

Grades of Discrimination

Indiscernibility, symmetry, and relativity

Tim Button
button@cantab.net
University of Cambridge

forthcoming in *Notre Dame Journal of Formal Logic*
final version submitted 26.iii.2015

1 Introduction

There are several relations which may fall short of genuine identity, but which behave like identity in important respects. Such *grades of discrimination* have recently been the subject of much philosophical and technical discussion.

Much of this discussion has been fuelled by considering *the Principle of the Identity of Indiscernibles*: the claim that indiscernible objects are always identical. The Principle is obviously of direct metaphysical interest (see [13]). But, within the philosophy of mathematics, the Principle has risen to prominence via the question of whether platonistically-minded structuralists can countenance structures with indiscernible but distinct positions (see [26, §1]). And, within the philosophy of physics, the central question has been whether quantum mechanics presented real-world counterexamples to the Principle (see [22]). As discussion has progressed, though, it has become increasingly clear that we must distinguish between different versions of ‘the’ Principle, corresponding to different notions of indiscernibility. This has spurred several philosophers to investigate the logical properties of these different notions (see [8], [17], [18]).

This paper completes that logical investigation. It exhaustively details, not just the properties of grades of indiscernibility, but the properties of all of the grades of discrimination. Indeed, this paper answers all of the mathematical questions that are natural at this level of abstraction.

There are three broad families of grades of discrimination. Grades of *indiscernibility* are defined in terms of satisfaction of certain first-order formulas, either with or without access to a primitive symbol that stands for genuine identity. They have been the focus of much recent philosophical attention. Grades of *symmetry* are defined in terms of isomorphisms. More specifically, they are defined in terms of symmetries (also known as automorphisms) on a structure. These grades have received some philosophical attention, though in a slightly less cohesive way than the grades of indiscernibility. Finally, grades of *relativity* are defined in terms of relateness correspondences, analogously to the grades of symmetry. The notion of a

relativeness correspondence has been studied by model-theorists, but is entirely absent from the philosophical literature on grades of discrimination. This paper rectifies this situation, introducing grades of relativity for the first time.

I mentioned, earlier, that the Principle of the Identity of Indiscernibles has been the main motivating force for interest in grades of discrimination. But it is now worth pausing to consider broader reasons for investigating the logical properties of the grades of discrimination.

The simplest reason to care about grades of discrimination is that they allow us to calibrate relationships of similarity and difference. More ambitiously, though, we might hope that some grade of discrimination will provide us with a genuinely illuminating answer to the question: When are objects identical? To take a simple example: set theory tells us that sets are identical iff they share all their members. Consequently, some grade of indiscernibility provides a suitable criterion of identity in set-theoretic contexts. To take a more contentious example: we might somehow become convinced that nature abhors a (non-trivial) symmetry. If so, then some grade of symmetry will provide a suitable criterion of identity in empirical contexts. The general hope, then, is that our grades of discrimination may furnish us with some non-trivial *criterion of identity* (in some context or other).

This search for a non-trivial criterion of identity need not be *reductive*. We might simply seek an illuminating constraint upon the conditions under which objects can be distinct. That said, some philosophers have hoped to find a reductive criterion for identity; that is, they have hoped to replace the identity primitive with some defined grade of discrimination. This reductive ambition is most prominent among those who have defended some Principle of Identity of Indiscernibles; such philosophers have therefore focussed on the various grades of indiscernibility. However, reductive ambitions might, in principle, be served equally well by considering either grades of symmetry or grades of relativity. (I revisit this in §8.)

Advancing a criterion of identity is not, however, simply a matter of selecting some grade of discrimination. As we shall see, each grade of discrimination is defined with respect to a (model-theoretic) signature. So, consider a signature which contains just a few monadic predicates which stand for eye colour. If we present a non-trivial criterion of identity, in the form of a grade of discrimination defined with respect to *this* signature, then we shall be forced to say, absurdly, that there is at most one person with brown eyes. Consequently, any philosopher who wants to advance a non-trivial criterion of identity must not only select some appropriate grade of discrimination, but must also stipulate the particular signature she has in mind. (I shall revisit this point several times below.)

In this paper, though, I am not aiming to *advance* any particular criterion of identity. My aim is only to provide a mathematical toolkit for anyone who

is interested in criteria of identity, whether reductive or non-reductive.

That toolkit is structured around three main results. Theorem 3.3 completely characterises the entailments between the grades of discrimination. Theorem 4.8 establishes a Galois Connection between isomorphisms and relativeness correspondences, which enables us better to understand the relationships between grades of symmetry and grades of relativity. And Theorem 6.12, is a Beth–Svenonius Theorem for logics without identity. By combining these three results, I answer several subsidiary questions concerning the grades of discrimination, including: which grades are equivalence relations (§5); which grades can be captured using sets of first-order formulas (§8); how they behave in finitary cases (§7); and how they behave in elementary extensions of structures (§6 and §9).

I now state some notational conventions.

I always use ‘ \mathcal{L} ’ to denote an arbitrary signature, i.e. a collection of constants, predicates and function-symbols. The philosophical discussion of grades of indiscernibility tends to be restricted to *relational* signatures, i.e. signatures which contain only predicates.¹ There are reasonable philosophical motivations for this: if we assume that each constant names exactly *one* object, then we seem to presuppose that we understand rather a lot about the notion of identity before we begin (see e.g. [2], [8, pp. 40–1]); more generally, the very idea of a *function* seems to presuppose the notion of identity; hence, if we want to avoid prejudging certain philosophical questions about identity, it might be wise to restrict our attention to relational signatures. The model-theoretic discussion of these issues is, though, less often restricted to relational signatures. There is a sensible technical motivation for this: many of the results hold in the more general case. Since this paper aims to provide philosophers with technical results, I shall allow signatures to contain both constants and functions, but I shall comment on the relational case when it is interestingly different. For technical ease, I treat constants as 0-place function-symbols.

Where \mathcal{L} is a signature, the \mathcal{L}^+ -formulas are the first-order formulas formed in the usual way using any \mathcal{L} -symbols and any symbols from standard first-order logic with identity. In particular, then, they may contain the symbol ‘=’, which always stands for (genuine) identity. The \mathcal{L}^- -formulas are those formed *without* using the symbol ‘=’. \mathcal{L}_n^+ is the set of \mathcal{L}^+ -formulas with free variables among ‘ v_1 ’, ..., ‘ v_n ’; similarly for \mathcal{L}_n^- .

I use swash fonts for structures and italic fonts for their associated domains. So, where \mathcal{M} is an \mathcal{L} -structure, its domain is M . Where $\bar{e} = \langle e_1, \dots, e_n \rangle$ and π is a function, $\pi(\bar{e}) = \langle \pi(e_1), \dots, \pi(e_n) \rangle$. Where Π is a two place relation, I write $\bar{d}\Pi\bar{e}$ to abbreviate $\langle d_1, e_1 \rangle, \dots, \langle d_n, e_n \rangle \in \Pi$.

¹An exception is Ladyman et al. [18, §6].

2 Twelve grades of discrimination

I start by defining six *grades of indiscernibility*; three grades of \mathcal{L}^- -indiscernibility, and three grades of \mathcal{L}^+ -indiscernibility.²

Definition 2.1. For any \mathcal{L} -structure \mathcal{M} with $a, b \in M$:

1. $a \approx_m b$ in \mathcal{M} iff $\mathcal{M} \models \phi(a) \leftrightarrow \phi(b)$ for all $\phi \in \mathcal{L}_1^-$
2. $a \approx_p b$ in \mathcal{M} iff $\mathcal{M} \models \phi(a, b) \leftrightarrow \phi(b, a)$ for all $\phi \in \mathcal{L}_2^-$
3. $a \approx b$ in \mathcal{M} iff $\mathcal{M} \models \phi(a, \bar{e}) \leftrightarrow \phi(b, \bar{e})$ for all $n < \omega$, all $\phi \in \mathcal{L}_{n+1}^-$ and all $\bar{e} \in M^n$

Similarly:

4. $a =_m b$ in \mathcal{M} iff $\mathcal{M} \models \phi(a) \leftrightarrow \phi(b)$ for all $\phi \in \mathcal{L}_1^+$
5. $a =_p b$ in \mathcal{M} iff $\mathcal{M} \models \phi(a, b) \leftrightarrow \phi(b, a)$ for all $\phi \in \mathcal{L}_2^+$
6. $a = b$ in \mathcal{M} iff a is identical to b

Here, ‘ $_p$ ’ indicates *pairwise* indiscernibility; ‘ $_m$ ’ indicates *monadic* indiscernibility; and no subscript indicates *complete* indiscernibility.

There are several alternative characterisations of \approx , two of which will prove useful (see Casanovas et al. [7, p. 508] and Ketland [17, Theorem 3.17]):

Lemma 2.2. For any \mathcal{L} -structure \mathcal{M} , the following are equivalent:

1. $a \approx b$ in \mathcal{M}
2. $\mathcal{M} \models \phi(a, \bar{e}) \leftrightarrow \phi(b, \bar{e})$ for all $n < \omega$, all atomic $\phi \in \mathcal{L}_{n+1}^-$ and all $\bar{e} \in M^n$
3. $\mathcal{M} \models \phi(a, a) \leftrightarrow \phi(a, b)$ for all $\phi \in \mathcal{L}_2^-$

Quine was the first philosopher to analyse all six grades of indiscernibility. His fullest discussion of them ended as follows:

May there even be many intermediate grades? The question is ill defined. By imposing special conditions on the form or content of the open sentence used in discriminating two objects, we could define any number of intermediate grades of discriminability, subject even to no linear order. What I have called moderate discriminability [i.e. $=_p$ or \approx_p], however, is the only intermediate grade that I see how to define at our present high level of generality. [25, p. 116]

Quine was right that Definition 2.1 essentially exhausts all of the grades of discrimination that are fairly natural, highly general, and which can be defined in terms of satisfaction of \mathcal{L}^- - and \mathcal{L}^+ -formulas.³ Nevertheless, other grades of discrimination are quite natural; we just need to consider

²This family of definitions has a long philosophical heritage, e.g.: Hilbert and Bernays [14, §5]; Quine [24, pp. 230–2], [25]; Caulton and Butterfield [8, §2.1, §3.2]; Ketland [16, pp. 306–7], [17, Definitions 2.3, 2.5]; Ladyman et al. [18, Definition 3.1, §6.4].

³Though Caulton and Butterfield [8] and Ladyman et al. [18] also explore the case of quantifier-free formulas.

alternative methods of definition. (The sense in which they are ‘intermediate’ grades will become clear in §3 and, as Quine conjectured, we will see that they are not linearly ordered.)

In particular, I shall introduce grades of discrimination that are defined in terms of *isomorphisms*. As a reminder:

Definition 2.3. Let \mathcal{M}, \mathcal{N} be \mathcal{L} -structures. An *isomorphism* from \mathcal{M} to \mathcal{N} is any bijection $\pi : M \rightarrow N$ such that:

1. $\bar{e} \in R^{\mathcal{M}}$ iff $\pi(\bar{e}) \in R^{\mathcal{N}}$, for all n -place \mathcal{L} -predicates R and all $\bar{e} \in M^n$
2. $\pi(f^{\mathcal{M}}(\bar{e})) = f^{\mathcal{N}}(\pi(\bar{e}))$, for all n -place \mathcal{L} -function-symbols f and all $\bar{e} \in M^n$

A *symmetry* on \mathcal{M} is an isomorphism from \mathcal{M} to \mathcal{M} .

Isomorphisms preserve \mathcal{L}^+ -formulas (see e.g. Marker [20, pp. 13–14]):

Lemma 2.4. Let \mathcal{M}, \mathcal{N} be \mathcal{L} -structures, and $\pi : \mathcal{M} \rightarrow \mathcal{N}$ be an isomorphism. For all $n < \omega$, all $\phi \in \mathcal{L}_n^+$ and all $\bar{e} \in M^n$:

$$\mathcal{M} \models \phi(\bar{e}) \text{ iff } \mathcal{N} \models \phi(\pi(\bar{e}))$$

There is therefore a good sense in which objects linked by a symmetry cannot be discriminated. Consequently, symmetries are a source of grades of discrimination, and I shall be interested in three distinct *grades of symmetry*:

Definition 2.5. For any \mathcal{L} -structure \mathcal{M} with $a, b \in M$:

1. $a \mathbf{s}_m b$ in \mathcal{M} iff there is a symmetry π on \mathcal{M} with $\pi(a) = b$
2. $a \mathbf{s}_p b$ in \mathcal{M} iff there is a symmetry π on \mathcal{M} with $\pi(a) = b$ and $\pi(b) = a$
3. $a \mathbf{s} b$ in \mathcal{M} iff there is a symmetry π on \mathcal{M} with $\pi(a) = b$, $\pi(b) = a$ and $\pi(x) = x$ for all $x \notin \{a, b\}$

These three grades have already received some philosophical attention;⁴ one of my aims is to incorporate them into the discussion in a systematic way.

In defining the notion of an isomorphism, the only object-language symbols which are mentioned are those of the signature; there is no need to mention ‘=’. Nevertheless, the notion of an isomorphism—and hence each grade of symmetry—straightforwardly depends upon the notion of identity. After all, an isomorphism is a bijection, which is to say it maps *unique* objects to *unique* objects, and *vice versa*. This dependence on identity is reflected in Lemma 2.4: symmetries preserve \mathcal{L}^+ -formulas.

If we want to avoid treating identity as a primitive—for philosophical or technical reasons—then the notion of an isomorphism is probably too strong.

⁴ \mathbf{s}_m is considered by Ketland [16], [17] under the name ‘structural indiscernibility’, and by Ladyman et al. [18] under the name ‘symmetry’. \mathbf{s}_p is considered by Ladyman et al. [18] under the name ‘full symmetry’. \mathbf{s} is considered by Ketland [16], [17], who writes ‘ π_{ab} ’ to indicate that $a \mathbf{s} b$.

In looking for a weaker notion, a first thought would be to consider functions between structures that need not be bijections. (In this regard, *strict homomorphisms* are sometimes considered.) But this is insufficiently concessive, since the very idea of a *function* presupposes the notion of identity, for a function maps each object (or n -tuple) to a *unique* object. Instead, then, we should consider structure-preserving *relations* that may hold between structures. The appropriate notion is provided by Casanovas et al. [7, Definition 2.5]; recall from §1 that $\bar{d}\Pi\bar{e}$ abbreviates $\langle d_1, e_1 \rangle, \dots, \langle d_n, e_n \rangle \in \Pi$:⁵

Definition 2.6. Let \mathcal{M}, \mathcal{N} be \mathcal{L} -structures. A *relativeness correspondence* from \mathcal{M} to \mathcal{N} is any relation $\Pi \subseteq M \times N$ with $\text{dom}(\Pi) = M$ and $\text{rng}(\Pi) = N$ such that:

1. $\bar{d} \in R^{\mathcal{M}}$ iff $\bar{e} \in R^{\mathcal{N}}$, for all n -place \mathcal{L} -predicates R and all $\bar{d}\Pi\bar{e}$
2. $f^{\mathcal{M}}(\bar{d})\Pi f^{\mathcal{N}}(\bar{e})$, for all n -place \mathcal{L} -function-symbols f and all $\bar{d}\Pi\bar{e}$

A *relativity* on \mathcal{M} is a relativeness correspondence from \mathcal{M} to \mathcal{M} .

Casanovas et al. [7, Proposition 2.6] show that relativeness correspondences preserve \mathcal{L}^- -formulas:

Lemma 2.7. Let \mathcal{M}, \mathcal{N} be \mathcal{L} -structures, and Π be a relativeness correspondence from \mathcal{M} to \mathcal{N} . For all $n < \omega$, all $\phi \in \mathcal{L}_n^-$, and all $\bar{d}\Pi\bar{e}$:

$$\mathcal{M} \models \phi(\bar{d}) \text{ iff } \mathcal{N} \models \phi(\bar{e})$$

There is therefore a good sense in which objects linked by a relativity cannot be discriminated. So, by simple analogy with the three grades of symmetry, I shall consider three *grades of relativity*:

Definition 2.8. For any \mathcal{L} -structure \mathcal{M} with $a, b \in M$:

1. $a \text{ r}_m b$ in \mathcal{M} iff there is a relativity Π on \mathcal{M} with $a\Pi b$
2. $a \text{ r}_p b$ in \mathcal{M} iff there is a relativity Π on \mathcal{M} with $a\Pi b$ and $b\Pi a$
3. $a \text{ r } b$ in \mathcal{M} iff there is a relativity Π on \mathcal{M} with $a\Pi b$, $b\Pi a$ and $x\Pi x$ for all $x \not\approx a$ and $x \not\approx b$

Unlike the grades of symmetry, the grades of relativity have not yet been considered by philosophers interested in grades of discrimination. However, there is no principled reason for this omission. Indeed, a central claim of this paper is that relativeness correspondences (and hence grades of relativity) are the appropriate \mathcal{L}^- -surrogate for isomorphisms (and hence grades of symmetry). This claim should already be plausible, given that we arrived at the notion of a relativeness correspondence by relaxing the notion of an isomorphism, and given the immediate comparison between Lemmas 2.4 and 2.7. The claim will receive further support during this paper.

⁵They credit a special case of this to Blok and Pigozzi [3, p. 343]; see also [28, §5].

For the reader's convenience, the following table summarises the twelve grades of discrimination:

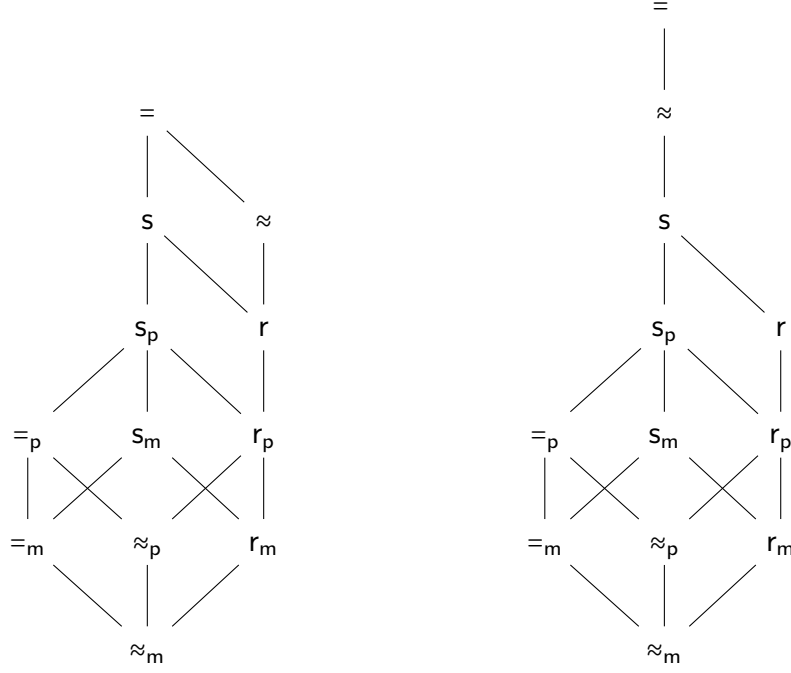
Grade	Informal gloss	Definition sketch
=	genuine identity	$a = b$
= _p	pairwise \mathcal{L}^+ -indiscernibility	$\phi(a, b) \leftrightarrow \phi(b, a)$, all $\phi \in \mathcal{L}_2^+$
= _m	monadic \mathcal{L}^+ -indiscernibility	$\phi(a) \leftrightarrow \phi(b)$, all $\phi \in \mathcal{L}_1^+$
\approx	complete \mathcal{L}^- -indiscernibility	$\phi(a, \bar{e}) \leftrightarrow \phi(b, \bar{e})$, all \bar{e} and $\phi \in \mathcal{L}_n^-$
\approx_p	pairwise \mathcal{L}^- -indiscernibility	$\phi(a, b) \leftrightarrow \phi(b, a)$, all $\phi \in \mathcal{L}_2^-$
\approx_m	monadic \mathcal{L}^- -indiscernibility	$\phi(a) \leftrightarrow \phi(b)$, all $\phi \in \mathcal{L}_1^-$
s	complete symmetry	a permutation $\pi(a) = b$, $\pi(b) = a$ and $\pi(x) = x$ all $x \notin \{a, b\}$
s _p	pairwise symmetry	a permutation $\pi(a) = b$, $\pi(b) = a$
s _m	monadic symmetry	a permutation $\pi(a) = b$
r	complete relativity	a relativity $a\Pi b$, $b\Pi a$ and $x\Pi x$ all $x \not\approx a$, $x \not\approx b$
r _p	pairwise relativity	a relativity $a\Pi b$, $b\Pi a$
r _m	monadic relativity	a relativity $a\Pi b$

3 Entailments between the grades

Having defined twelve grades of discrimination, my first task is to characterise the relationships between them. More precisely: I shall build upon some existing results (mentioned in endnotes) to provide a complete account of the entailments and non-entailments between the various grades of discrimination.

For any two grades of discrimination R and S, say that R *entails* S *iff* for any structure \mathcal{M} and any $a, b \in M$, if aRb then aSb in \mathcal{M} . Entailment is relativised to particular classes of structures—e.g. to structures with relational signatures—in the obvious way. In §7, I shall consider the special case of entailments where we restrict our attention to *finite* structures. However, the target result for this section is the general case:

Theorem 3.3 (Entailments between the grades). These Hasse Diagrams characterise the entailments between our grades of discrimination:



The left diagram considers the case of arbitrary signatures; the right diagram considers entailment when restricted to structures with relational signatures.

To explain the notation: there is a path down the page from R to S *iff* R entails S . So in the case of arbitrary signatures: $=$ entails \approx ; \approx does not entail $=$; s does not entail \approx ; and \approx does not entail s . In the case of relational signatures: \approx entails s ; hence \approx entails s_p ; etc. I shall start by proving the entailments:⁶

Lemma 3.1. For structures with arbitrary signatures:

1. $=$ entails both \approx and s
2. \approx entails r
3. $=_p$ entails \approx_p , and $=_m$ entails \approx_m
4. $=_p$ entails $=_m$, and \approx_p entails \approx_m
5. s_p entails $=_p$, and s_m entails $=_m$
6. r_p entails \approx_p , and r_m entails \approx_m
7. s entails r , s_p entails r_p , and s_m entails r_m
8. s entails s_p , and s_p entails s_m
9. r entails r_p , and r_p entails r_m

For structures with relational signatures, but not in general:

10. \approx entails s

⁶Quine [25] proves case (4); see also Ketland [16, p. 307], [17, §3.2], Ladyman et al. [18, Theorem 5.2]. Caulton and Butterfield [8, Theorem 1] prove case (5) when restricted to relational signatures; see also Ketland [17, Lemma 3.22] and Ladyman et al. [18, Theorem 9.17, 9.20]. Ketland [17, Theorem 3.23] proves case (10).

Proof. (1). The identity map is a symmetry.

(2). The relation given by $x \Pi x$ iff $x \approx x$ is a relativity.

(3)–(4). Immediate from the definitions.

(5)–(6). Immediate from Lemmas 2.4 and 2.7.

(7). Every symmetry can be regarded as a relativity.

(8)–(9). Immediate from the definitions.

(10). If $a \approx b$, then for any $n < \omega$, any atomic formula $\phi \in \mathcal{L}_n^-$, and any $\bar{e} \in (M \setminus \{a, b\})^n$, we have $\mathcal{M} \models \phi(a, \bar{e}) \leftrightarrow \phi(b, \bar{e})$. \square

It remains to demonstrate the non-entailments:⁷

Lemma 3.2. For structures with relational signatures:

1. \approx does not entail $=$
2. s does not entail \approx
3. r does not entail $=_m$
4. s_m does not entail \approx_p
5. s_p does not entail r
6. $=_p$ does not entail r_m

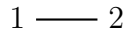
Moreover, for structures with arbitrary signatures:

7. \approx does not entail $=_m$

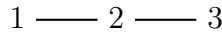
Proof. (1). In this unlabelled graph, \mathcal{A} , we have $1 \approx 2$ but $1 \neq 2$:



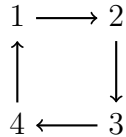
(2). In this unlabelled graph, \mathcal{B} , we have $1 s 2$ but $1 \not\approx 2$:



(3). In this unlabelled graph, \mathcal{C} , we have $1 r 2$ but $1 \neq_m 2$:



(4). In this unlabelled directed graph, \mathcal{D} , we have $1 s_m 2$ but $1 \not\approx_p 2$:



⁷Ketland [17, p. 8] and Ladyman et al. [18, Theorem 5.2] use \mathcal{A} to prove case (1); see also Button [5, p. 218] and Ketland [16, p. 309]. Ladyman et al. [18, Theorem 9.17] use \mathcal{B} to prove case (2), noting that it is the analogue of Black's [2] two-sphere world. Button [5, p. 218], Ketland [16, p. 310] and Ladyman et al. [18, Theorem 7.12] use an example like \mathcal{C} . Ladyman et al. [18, Theorem 5.2] use an example like \mathcal{D} . Caulton and Butterfield [8, pp. 60–2] and Ladyman et al. [18, Theorem 9.20] use examples like \mathcal{E} .

- (5). In \mathcal{D} , again, we have $1 \text{ s}_p 3$ but $1 \not\text{f}_3 3$.
(6). Let \mathcal{E} be the disjoint union of a complete countably-infinite graph with a complete uncountable graph, i.e.:

$$E := \mathbb{R} \\ R^{\mathcal{E}} := \{\langle n, m \rangle \in \mathbb{N}^2 \mid n \neq m\} \cup \{\langle p, q \rangle \in (\mathbb{R} \setminus \mathbb{N})^2 \mid p \neq q\}$$

By taking a Skolem Hull of \mathcal{E} containing $1 \in \mathbb{N}$ and any $e \in \mathbb{R} \setminus \mathbb{N}$, we see that $1 =_p e$. Now suppose that Π is a relativity with $1\Pi e$. Since Π must preserve the edges of the graph, and every element in either ‘cluster’ has an edge to every element in the cluster except itself, Π must be a bijection between \mathbb{N} and $\mathbb{R} \setminus \mathbb{N}$. Contradiction; so $1 \not\text{f}_m e$.

- (7). Augment \mathcal{A} by adding a single constant which picks out 1. □

It is simple to check that Lemmas 3.1 and 3.2 yield Theorem 3.3. This Theorem allows us to compare the consequences of imposing various grades of discrimination as criteria of identity.

I should comment briefly on the philosophical significance of the constructions used in Lemma 3.2. The existence of \mathcal{A} is guaranteed by absolutely standard model theory. However, \mathcal{A} contains two distinct objects that are ‘blank’: from the perspective of \mathcal{A} , these objects have no properties or relations to anything, so that their distinctness must be *brute*. And this might suggest that the use of absolutely standard model theory begs the question against anyone who believes in a non-trivial criterion of identity. Fortunately it does not, but it is worth carefully explaining why.

Let Fran be a philosopher who advocates a non-trivial criterion of identity: in particular, Fran thinks that x and y are identical iff $x \approx y$. However, bearing in mind the discussion of §1—particularly of a signature which allows us only to describe eye colour—Fran advances this criterion of identity with respect to some *particular* signature, \mathcal{F} . Now, if \mathcal{A} is presented as an \mathcal{F} -structure, then Fran will certainly deny that \mathcal{A} could exist. However, Fran can make sense of \mathcal{A} by regarding it as a \mathcal{G} -structure, where \mathcal{G} is a signature which is impoverished compared with \mathcal{F} . Construed thus, \mathcal{A} begs no question against Fran, because it poses no threat to her proposed criterion of identity.

To be clear: I am not trying to endorse or defend Fran’s position.⁸ My point is simply that everyone, including Fran, can make sense of standard model theory.

4 A Galois Connection

Theorem 3.3 graphically demonstrates that grades of symmetry are to grades of \mathcal{L}^+ -indiscernibility as grades of relativity are to grades of \mathcal{L}^- -indiscernibility.

⁸I was once on Fran’s side [5, p. 220]; but I have changed my mind [6, p. 211 n. 8].

In this section, I develop this point by outlining a Galois Connection between isomorphisms and relativeness correspondences. (The results of this section can be fruitfully compared with those of [Bonney and Engström](#) [4]; we discovered our results independently.)

Lemma 2.4 has an obvious converse: every bijective map which preserves all \mathcal{L}^+ -formulas is an isomorphism. However, there is no converse to Lemma 2.7. To make this more precise, consider the following definition:

Definition 4.1. Let \mathcal{M}, \mathcal{N} be \mathcal{L} -structures. A *near-correspondence* from \mathcal{M} to \mathcal{N} is any relation $\Pi \subseteq M \times N$ with $\text{dom}(\Pi) = M$ and $\text{rng}(\Pi) = N$ such that, for all $n < \omega$, all $\phi \in \mathcal{L}_n^-$, and all $\bar{d}\Pi\bar{e}$:

$$\mathcal{M} \models \phi(\bar{d}) \text{ iff } \mathcal{N} \models \phi(\bar{e})$$

Lemma 2.7 states that every relativeness correspondence is a near-correspondence. But the converse fails. Let \mathcal{F} be an $\{f\}$ -structure, defined as follows:

$$\begin{aligned} F &= \{1, 2\} \\ f^{\mathcal{F}}(1) &= f^{\mathcal{F}}(2) = 2 \end{aligned}$$

Then $\Pi = \{\langle 1, 2 \rangle, \langle 2, 1 \rangle\}$ is clearly a near-correspondence from \mathcal{F} to \mathcal{F} , but not a relativeness correspondence.

However, there is an elegant connection between near-correspondences (and hence relativeness correspondences) and isomorphisms on the models we obtain by quotienting using \approx . The use of such quotients is standard in model theory without identity, and the central idea is summed up in the following Definition and Lemma (see [Casanovas et al.](#) [7, Definition 2.3–2.4]):⁹

Definition 4.2. Let \mathcal{M} be any \mathcal{L} -structure. Then $\underline{\mathcal{M}}$ is the \mathcal{L} -structure obtained by quotienting \mathcal{M} by \approx . We denote its members with $\underline{a}_{\mathcal{M}} = \{b \in M \mid a \approx b \text{ in } \mathcal{M}\}$ and, when no confusion can arise, we dispense with the subscript, talking of \underline{a} rather than $\underline{a}_{\mathcal{M}}$. Now $\underline{\mathcal{M}}$ is defined as follows:

$$\begin{aligned} \underline{M} &= \{\underline{a} \mid a \in M\} \\ R^{\underline{\mathcal{M}}} &= \{\underline{\bar{e}} \in \underline{M}^n \mid \bar{e} \in R^{\mathcal{M}}\} \quad \text{all } n\text{-place } \mathcal{L}\text{-predicates } R \\ f^{\underline{\mathcal{M}}}(\underline{\bar{e}}) &= \underline{f^{\mathcal{M}}(\bar{e})} \quad \text{all } n\text{-place } \mathcal{L}\text{-function-symbols } f \text{ and all } \bar{e} \in M^n \end{aligned}$$

Lemma 4.3. Let \mathcal{M} be an \mathcal{L} -structure. For all $n < \omega$, all $\phi \in \mathcal{L}_n^-$ and all $\bar{e} \in M^n$:

$$\mathcal{M} \models \phi(\bar{e}) \text{ iff } \underline{\mathcal{M}} \models \phi(\underline{\bar{e}})$$

⁹[Casanovas et al.](#) trace the definition and lemma back to [Monk](#) [21, presumably Exercises 29.33–34]. This has recently been rediscovered by philosophers, e.g. Ketland [16, p. 307 n. 10], [17, Theorem 3.12].

Casanovas et al. [7, Proposition 2.6] note that there is a relativeness correspondence from \mathcal{M} to \mathcal{N} iff $\underline{\mathcal{M}} \cong \underline{\mathcal{N}}$. I wish to build on this; and I begin with some definitions:

Definition 4.4. For any \mathcal{L} -structures \mathcal{M}, \mathcal{N} :

1. $\mathbf{N}(\mathcal{M}, \mathcal{N})$ is the set of near-correspondences from \mathcal{M} to \mathcal{N}
2. $\mathbf{I}(\mathcal{M}, \mathcal{N})$ is the set of isomorphisms from $\underline{\mathcal{M}}$ to $\underline{\mathcal{N}}$
3. $\mathbf{e}(\mathcal{M}, \mathcal{N}) : \mathbf{I}(\mathcal{M}, \mathcal{N}) \rightarrow \mathbf{N}(\mathcal{M}, \mathcal{N})$ is given by: $a\pi^{\mathbf{e}}b$ iff $\pi(\underline{a}) = \underline{b}$
4. $\mathbf{c}(\mathcal{M}, \mathcal{N}) : \mathbf{N}(\mathcal{M}, \mathcal{N}) \rightarrow \mathbf{I}(\mathcal{M}, \mathcal{N})$ is given by: $\Pi^{\mathbf{c}}(\underline{a}) = \underline{b}$ iff there are $a' \approx a$ and $b' \approx b$ such that $a'\Pi b'$

Say that $\Pi \in \mathbf{N}(\mathcal{M}, \mathcal{N})$ is *maximal* iff no strict superset of Π is in $\mathbf{N}(\mathcal{M}, \mathcal{N})$.

I prove that these are genuine definitions, i.e. that $\mathbf{e}(\mathcal{M}, \mathcal{N})$ and $\mathbf{c}(\mathcal{M}, \mathcal{N})$ are functions. I begin with $\mathbf{e}(\mathcal{M}, \mathcal{N})$:

Lemma 4.5. If $\pi \in \mathbf{I}(\mathcal{M}, \mathcal{N})$, then $\pi^{\mathbf{e}}$ is a relativeness correspondence, and hence $\pi^{\mathbf{e}} \in \mathbf{N}(\mathcal{M}, \mathcal{N})$.

Proof. Fix $n < \omega$ and suppose that $\bar{d}\pi^{\mathbf{e}}\bar{e}$; i.e. that $\pi(\bar{d}) = \bar{e}$. For each n -place L -predicate R , observe:

$$\bar{d} \in R^{\mathcal{M}} \text{ iff } \bar{d} \in R^{\underline{\mathcal{M}}} \text{ iff } \pi(\bar{d}) \in R^{\underline{\mathcal{N}}} \text{ iff } \bar{e} \in R^{\underline{\mathcal{N}}} \text{ iff } \bar{e} \in R^{\mathcal{N}}$$

For each n -place L -function-symbol f , observe:

$$\pi(\underline{f^{\mathcal{M}}(\bar{d})}) = \pi(f^{\underline{\mathcal{M}}}(\bar{d})) = f^{\underline{\mathcal{N}}}(\pi(\bar{d})) = f^{\underline{\mathcal{N}}}(\bar{e}) = \underline{f^{\mathcal{N}}(\bar{e})}$$

so that $f^{\mathcal{M}}(\bar{d})\pi^{\mathbf{e}}f^{\mathcal{N}}(\bar{e})$. Hence $\pi^{\mathbf{e}}$ is a relativeness correspondence, and so a near-correspondence by Lemma 2.7. \square

To show that $\mathbf{c}(\mathcal{M}, \mathcal{N})$ is a function, we need a subsidiary result:

Lemma 4.6. Let Π be a near-correspondence from \mathcal{M} to \mathcal{N} , with $a\Pi b$ and $a'\Pi b'$. Then $a \approx a'$ in \mathcal{M} iff $b \approx b'$ in \mathcal{N} .

Proof. Suppose $a \approx a'$ in \mathcal{M} ; then using Lemmas 2.7 and 2.2:

$$\mathcal{N} \models \phi(b, b) \text{ iff } \mathcal{M} \models \phi(a, a) \text{ iff } \mathcal{M} \models \phi(a, a') \text{ iff } \mathcal{N} \models \phi(b, b')$$

So $b \approx b'$ in \mathcal{N} , by Lemma 2.2. The converse is similar. \square

It follows that $\mathbf{c}(\mathcal{M}, \mathcal{N})$ is a function:

Lemma 4.7. If $\Pi \in \mathbf{N}(\mathcal{M}, \mathcal{N})$, then $\Pi^{\mathbf{c}} \in \mathbf{I}(\mathcal{M}, \mathcal{N})$.

Proof. Lemma 4.6 immediately yields that $\Pi^{\mathbf{c}}(\bar{a})$ is a well-defined function, and indeed an injection. $\Pi^{\mathbf{c}}$ is a surjection, since $\text{rng}(\Pi) = N$. It remains to

show that Π^c preserves structure. For the remainder of the proof, fix $n < \omega$ and $\bar{a}, \bar{b} \in M^n$ such that $\Pi^c(\bar{a}) = \bar{b}$.

Let R be any n -place \mathcal{L} -predicate. For each $1 \leq i \leq n$ we have $a'_i \approx a_i$ and $b'_i \approx b_i$ such that $a'_i \Pi b'_i$; and hence:

$$\bar{a} \in R^{\mathcal{M}} \text{ iff } \bar{a}' \in R^{\mathcal{M}} \text{ iff } \bar{b}' \in R^{\mathcal{N}} \text{ iff } \bar{b} \in R^{\mathcal{N}} \text{ iff } \Pi^c(\bar{a}) \in R^{\mathcal{N}}$$

Let f be any n -place \mathcal{L} -function-symbol. For all $\phi(v_1, v_2) \in \mathcal{L}_2^-$, define $\phi_f(v_1, \bar{x})$ as $\phi(v_1, f(\bar{x}))$; and let $b' \in N$ be such that $f^{\mathcal{M}}(\bar{a}) \Pi b'$. Invoking Lemma 2.2:

$$\begin{aligned} \mathcal{N} \models \phi(b', b') &\text{ iff } \mathcal{M} \models \phi(f^{\mathcal{M}}(\bar{a}), f^{\mathcal{M}}(\bar{a})) \\ &\text{ iff } \mathcal{M} \models \phi_f(f^{\mathcal{M}}(\bar{a}), \bar{a}) \\ &\text{ iff } \mathcal{N} \models \phi_f(b', \bar{b}) \\ &\text{ iff } \mathcal{N} \models \phi(b', f^{\mathcal{N}}(\bar{b})) \end{aligned}$$

Hence $b' \approx f^{\mathcal{N}}(\bar{b})$ by Lemma 2.7. Now:

$$\Pi^c(f^{\mathcal{M}}(\bar{a})) = \Pi^c(\underline{f^{\mathcal{M}}(\bar{a})}) = \underline{b'} = \underline{f^{\mathcal{N}}(\bar{b})} = f^{\mathcal{N}}(\bar{b}) = f^{\mathcal{N}}(\Pi^c(\bar{a}))$$

so that functions are preserved. \square

Lemmas 4.5 and 4.7 together show that Definition 4.4 is a proper definition. Its significance resides in the following:

Theorem 4.8 (Galois Connection on \approx -quotients). For each $\Pi \in \mathbf{N}(\mathcal{M}, \mathcal{N})$ and each $\pi \in \mathbf{I}(\mathcal{M}, \mathcal{N})$: $\Pi^c = \pi$ iff $\Pi \subseteq \pi^e$.

Proof. Left-to-right. Suppose $\Pi^c = \pi$. Fix $\langle d, e \rangle \in \Pi$; then $\pi(\underline{d}) = \underline{e}$, so $d \pi^e e$.

Right-to-left. Suppose $\Pi \subseteq \pi^e$. Where $\Pi^c(\underline{d}) = \underline{e}$, there are $d' \approx d$ and $e' \approx e$ such that $d' \Pi e'$ and hence $d' \pi^e e'$; so $\underline{e} = \underline{e'} = \pi(\underline{d'}) = \pi(\underline{d})$. Hence $\Pi^c(\underline{d}) = \pi(\underline{d})$, for all $\underline{d} \in \underline{M}$. \square

This Theorem highlights the depth of the connection between isomorphisms and relativity correspondences. Additionally, it strengthens the claim that relativity correspondences are the \mathcal{L}^- -analogue of isomorphisms. For, given that there are near-correspondences that are not relativity correspondences, one might have worried that relativity correspondences *compete* with the near-correspondences to be the \mathcal{L}^- -analogue of isomorphism. However, the appearance of competition vanishes, once we consider some consequences of the Galois Connection:

Lemma 4.9. For any \mathcal{L} -structures \mathcal{M}, \mathcal{N} :

1. $\mathbf{c}(\mathcal{M}, \mathcal{N}) \circ \mathbf{e}(\mathcal{M}, \mathcal{N})$ is the identity function
2. $\mathbf{e}(\mathcal{M}, \mathcal{N}) \circ \mathbf{c}(\mathcal{M}, \mathcal{N})$ is idempotent

3. If $\pi \in \mathbf{I}(\mathcal{M}, \mathcal{N})$, then π^e is maximal
4. If $\Pi \in \mathbf{N}(\mathcal{M}, \mathcal{N})$, then $(\Pi^c)^e$ is the unique maximal relativeness correspondence that extends Π

Proof. (1)–(2). Immediate from the fact that this is a Galois Connection with the partial-ordering on $\mathbf{I}(\mathcal{M}, \mathcal{N})$ being identity.

(3). Let $\Sigma \in \mathbf{I}(\mathcal{M}, \mathcal{N})$ be such that $\pi^e \subseteq \Sigma$, and suppose $a\Sigma b$. By Lemma 4.7, $\Sigma^c \in \mathbf{N}(\mathcal{M}, \mathcal{N})$, with $\Sigma^c(\underline{a}) = \underline{b}$. For any d such that $a\pi^e d$, we have $a\Sigma d$, and hence $\Sigma^c(\underline{a}) = \underline{d}$, so that $b \approx d$. Hence $\pi(\underline{a}) = \underline{d} = \underline{b}$, and so $a\pi^e b$.

(4). Lemma 4.5, our Galois Connection, and case (3) show that $(\Pi^c)^e$ is a maximal relativeness correspondence extending Π . To show uniqueness, let Σ be any maximal relativeness correspondence extending Π . Consider any $a, b \in M$ such that $a(\Pi^c)^e b$. Then there are $a' \approx a$, $b' \approx b$ such that $a'\Pi b'$, and hence such that $a'\Sigma b'$, since $\Pi \subseteq \Sigma$. Hence, for any $\bar{d}\Sigma\bar{e}$, and any $\phi \in \mathcal{L}_{n+1}^-$, by Lemma 2.7:

$$\mathcal{M} \models \phi(a, \bar{d}) \text{ iff } \mathcal{M} \models \phi(a', \bar{d}) \text{ iff } \mathcal{N} \models \phi(b', \bar{e}) \text{ iff } \mathcal{N} \models \phi(b, \bar{e})$$

Consequently, $\Theta = \Sigma \cup \{\langle a, b \rangle\}$ is a near-correspondence. So $(\Theta^c)^e$ is a maximal relativeness correspondence extending Σ ; but Σ is itself maximal, so $a\Sigma b$. Generalising, $(\Pi^c)^e \subseteq \Sigma$. Since $(\Pi^c)^e$ is maximal, $(\Pi^c)^e = \Sigma$. \square

The preceding result tells us that every near-correspondence expands to a relativeness correspondence. Accordingly, there is no genuine competition between near-correspondences and relativeness correspondences. Indeed, instead of defining the grades of relativity in terms of relativeness correspondences, we could have defined them in terms of near-correspondences. Or, even more simply, we could have defined them in terms of symmetries on quotient models, as shown by the following immediate consequence of the preceding results:

Lemma 4.10. For any \mathcal{L} -structure \mathcal{M} :

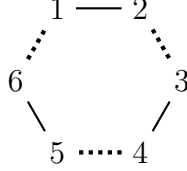
1. $a r_m b$ in \mathcal{M} iff $\underline{a} s_m \underline{b}$ in $\underline{\mathcal{M}}$
2. $a r_p b$ in \mathcal{M} iff $\underline{a} s_p \underline{b}$ in $\underline{\mathcal{M}}$
3. $a r b$ in \mathcal{M} iff $\underline{a} s \underline{b}$ in $\underline{\mathcal{M}}$

5 Equivalence relations

At various points, I have described the grades of discrimination as behaving like identity. A natural question is whether the grades of discrimination behave like identity in being *equivalence* relations. (Note that I implicitly relied upon the fact that \approx is an equivalence relation in defining the \approx -quotient structure.) Ladyman et al. [18, Theorem 10.22] have partially answered this question, in noting that $=_p$ and \approx_p are not transitive (in general). The following result, which employs our Galois Connection, completes the picture.

Theorem 5.1. $=_p$, \approx_p , s_p and r_p are reflexive and symmetric, but not transitive (in general); the remaining eight grades of discrimination are equivalence relations.

Proof. Consider the following coloured graph, \mathcal{G} :



Here, $1 s_p 2$ and $2 s_p 3$, but $1 \not\approx_p 3$. By Theorem 3.3, this establishes that $=_p$, \approx_p , s_p and r_p are not transitive (in general).

The reflexivity and symmetry of all the grades of indiscernibility are immediate from their definitions, as is the transitivity of $=$, \approx , $=_m$ and \approx_m .

It is routine to check that all three grades of symmetry are symmetric and reflexive, and that s and s_m are transitive. Lemma 4.10 entails that the same is true for the respective grades of relativity. \square

Since identity is surely transitive, Theorem 5.1 might seem to provide a knockdown argument against treating any of $=_p$, \approx_p , s_p and r_p as a criterion of identity. However, this point is a little more subtle than it might initially seem.

Consider the discussion of \mathcal{A} , at the end of §3. \mathcal{A} might have seemed to present a counterexample to treating \approx as a criterion of identity. But any philosopher who advocates such a criterion, such as Fran, will maintain that we can make sense of \mathcal{A} by (and only by) treating it as a structure of some artificially *restricted* signature. At that point, \mathcal{A} no longer presents a counterexample to Fran’s proposed criterion of identity, which she advances with respect to some *richer* signature.

With this in mind, consider Rach, a philosopher who advocates r_p as a criterion of identity. \mathcal{G} might seem to pose problems for Rach. But if \mathcal{G} is presented with regard to Rach’s preferred signature, then it violates her proposed criterion of identity even before we consider issues about transitivity: after all, \mathcal{G} is to contain objects which are distinct but (‘genuinely’) pairwise symmetric. Accordingly, Rach will maintain that we can make sense of \mathcal{G} by (and only by) treating it as a structure of some artificially *restricted* signature. And at that point, \mathcal{G} no longer demonstrates the non-transitivity of Rach’s proposed criterion of identity, which she advances with respect to some *richer* signature.

The situation, then, is slightly odd. From the perspective of anyone who thinks that identity is more fine-grained than any of $=_p$, \approx_p , s_p and r_p , these four grades of discrimination fail to behave like identity in an absolutely crucial sense, in failing to be transitive. (This is why Ladyman et al. [18,

p. 23] suggest that $=_p$ and \approx_p violate a plausible ‘minimal requirement’ on any notion of indiscernibility.) But it does not immediately follow that one cannot *propose* one of these four grades as a criterion of identity.

6 Connections to definability theory

I now want to explore some natural technical questions which have not featured on the radar of philosophers interested in grades of discrimination. These questions concern the relationship between grades of discrimination and *elementary extensions*, and they relate to definability theory. My answers to these questions, together with the Galois Connection of §4, will yield interesting entailments between the different grades in special cases (to be discussed in §7). To be clear on terminology:

Definition 6.1. Let \mathcal{M} and \mathcal{N} be \mathcal{L} -structures. Say that $\mathcal{M} <^+ \mathcal{N}$ iff for all $n < \omega$, all $\phi \in \mathcal{L}_n^+$, and all $\bar{e} \in M^n$:

$$\mathcal{M} \models \phi(\bar{e}) \text{ iff } \mathcal{N} \models \phi(\bar{e})$$

Say that $\mathcal{M} <^- \mathcal{N}$ iff the above holds with \mathcal{L}_n^- in place of \mathcal{L}_n^+ .

There is a classic result connecting \mathcal{L}^+ -indiscernibility with the existence of a symmetry in some elementary extension (see e.g. Marker [20, Proposition 4.1.5]):

Theorem 6.2. For any \mathcal{L} -structure \mathcal{M} , the following are equivalent:

1. $\mathcal{M} \models \phi(\bar{a}) \leftrightarrow \phi(\bar{b})$, for all $\phi \in \mathcal{L}_n^+$
2. There is an $\mathcal{N} >^+ \mathcal{M}$ and a symmetry π on \mathcal{N} such that $\pi(\bar{a}) = \bar{b}$

For present purposes, the immediate import of Theorem 6.2 is that it yields a new way to characterise $=_p$ and $=_m$:

Lemma 6.3. For any \mathcal{L} -structure \mathcal{M} :

1. $a =_m b$ in \mathcal{M} iff there is an $\mathcal{N} >^+ \mathcal{M}$ in which $a s_m b$
2. $a =_p b$ in \mathcal{M} iff there is an $\mathcal{N} >^+ \mathcal{M}$ in which $a s_p b$

This raises a natural question: Is there an \mathcal{L}^- -analogue of Theorem 6.2? There certainly is; but to show this, I need two definitions. First, I need the ordinary notion of a diagram:¹⁰

Definition 6.4. Let \mathcal{L} be any signature and X be any set. Then $\mathcal{L}(X)$ is the signature formed by augmenting \mathcal{L} with each member of X as a (new)

¹⁰Whilst this notion of diagram invokes ‘=’, Dellunde shows that there is a perfectly workable notion of diagram which does not employ ‘=’.

constant. Where \mathcal{M} is an \mathcal{L} -structure, $\text{Diag}^+(\mathcal{M})$ is the set of $\mathcal{L}^+(M)$ -sentences satisfied by the $\mathcal{L}(M)$ -structure formed by letting each $e \in M$ name itself.

Next, I need the \mathcal{L}^- -analogue for a partial elementary map:

Definition 6.5. Let \mathcal{M}, \mathcal{N} be \mathcal{L} -structures. A *proto-correspondence from \mathcal{M} to \mathcal{N}* is any relation $\Pi \subseteq M \times N$ such that, for all $n < \omega$, all $\phi \in \mathcal{L}_n^-$, and all $\bar{a}\Pi\bar{e}$:

$$\mathcal{M} \models \phi(\bar{a}) \text{ iff } \mathcal{N} \models \phi(\bar{e})$$

So a near-correspondence from \mathcal{M} to \mathcal{N} is a proto-correspondence with domain M and range N . The proof of the \mathcal{L}^- -analogue of Theorem 6.2 now amounts to little more than a tweak to Marker's proof of Theorem 6.2.¹¹ I start with two type-realising constructions:

Lemma 6.6. Let Π be a proto-correspondence from \mathcal{M} to \mathcal{N} . For any $a \in M$, there is as an $\mathcal{O} >^+ \mathcal{N}$ with some $b \in O$ such that $\Pi \cup \{\langle a, b \rangle\}$ is a proto-correspondence from \mathcal{M} to \mathcal{O} .

Proof. Define:

$$\Gamma = \{\phi(v, \bar{e}) \in \mathcal{L}_1^-(\text{rng}(\Pi)) \mid \text{for some } n < \omega, \text{ some } \phi \in \mathcal{L}_{n+1}^-, \text{ and some } \bar{a}\Pi\bar{e}, \\ \text{we have } \mathcal{M} \models \phi(a, \bar{d})\}$$

Consider any $\phi(v, \bar{e}) \in \Gamma$; since $\mathcal{M} \models \exists v \phi(v, \bar{d})$ and Π is a proto-correspondence, $\mathcal{N} \models \exists v \phi(v, \bar{e})$. Equally, \mathcal{N} can be treated as a model of $\text{Diag}^+(\mathcal{N})$. So any finite subset of $\Gamma \cup \text{Diag}^+(\mathcal{N})$ is satisfiable. Hence, by Compactness, there is a model of $\Gamma \cup \text{Diag}^+(\mathcal{N})$, which we can regard as $\mathcal{O} >^+ \mathcal{N}$. Now simply let $\Sigma = \Pi \cup \{\langle a, b \rangle\}$, where $\mathcal{O} \models \Gamma(b)$. \square

Lemma 6.7. Let Π be a proto-correspondence from \mathcal{M} to \mathcal{N} with $\mathcal{M} <^+ \mathcal{N}$. Then there is some $\mathcal{O} >^+ \mathcal{N}$ and a proto-correspondence $\Sigma \supseteq \Pi^{-1}$ from \mathcal{N} to \mathcal{O} with $\text{dom}(\Sigma) = N$.

Proof. We construct an elementary chain. Since Π is a proto-correspondence and $\mathcal{M} <^+ \mathcal{N}$, we have that for all $\phi \in \mathcal{L}_n^-$ and all $\bar{a}\Pi\bar{b}$:

$$\mathcal{N} \models \phi(\bar{b}) \text{ iff } \mathcal{M} \models \phi(\bar{a}) \text{ iff } \mathcal{N} \models \phi(\bar{a})$$

¹¹In more detail: my Lemma 6.6 tweaks Marker's Lemma 4.16; my Lemma 6.7 tweaks Marker's Corollary 4.1.7 (cf. also Casanovas et al.'s [7] Lemma 2.7); and my Lemma 6.8 tweaks Marker's Proposition 4.1.5. The main difference is that I use proto-correspondences rather than partial elementary maps, and in the final step I require a detour, via Lemma 4.9, to obtain a relativity.

Defining $\mathcal{O}_0 = \mathcal{N}$ and $\Sigma_0 = \Pi^{-1}$, observe that Σ_0 is a proto-correspondence from \mathcal{N} to \mathcal{O}_0 . This is our initial stage in the chain.

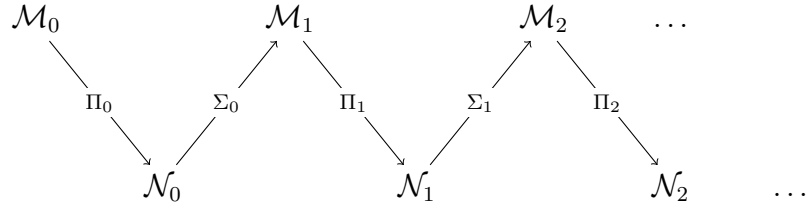
Now let $\{e_\alpha \mid \alpha < \kappa\}$ exhaustively enumerate N , let $D_\alpha = \text{dom}(\Sigma_0) \cup \{e_\beta \mid \beta < \alpha\}$ for each $\alpha < \kappa$, and proceed recursively:

- Stage $\alpha + 1$: Given a proto-correspondence Σ_α from \mathcal{N} to \mathcal{O}_α with $\text{dom}(\Sigma_\alpha) = D_\alpha$, use Lemma 6.6 to obtain an $\mathcal{O}_{\alpha+1} >^+ \mathcal{O}_\alpha$ and a proto-correspondence $\Sigma_{\alpha+1} \supseteq \Sigma_\alpha$ from \mathcal{N} to $\mathcal{O}_{\alpha+1}$ with $\text{dom}(\Sigma_{\alpha+1}) = D_{\alpha+1}$.
- Stage α , with α a limit ordinal: let $\mathcal{O}_\alpha = \bigcup_{\beta < \alpha} \mathcal{O}_\beta$ and $\Sigma_\alpha = \bigcup_{\beta < \alpha} \Sigma_\beta$.

Now let $\mathcal{O} = \bigcup_{\alpha < \kappa} \mathcal{O}_\alpha$ and $\Sigma = \bigcup_{\alpha < \kappa} \Sigma_\alpha$. \square

Lemma 6.8. Let \mathcal{M} be an \mathcal{L} -structure with $\bar{a}, \bar{b} \in M^n$ such that $\mathcal{M} \models \phi(\bar{a}) \leftrightarrow \phi(\bar{b})$ for all $\phi \in \mathcal{L}_n^-$. Then there is some $\mathcal{N} >^+ \mathcal{M}$ and a near-correspondence Π from \mathcal{N} to \mathcal{N} such that $\bar{a}\Pi\bar{b}$.

Proof. Given $\mathcal{M}, \bar{a}, \bar{b}$ as described, we have a proto-correspondence Π_0 from \mathcal{M} to \mathcal{M} with $\bar{a}\Pi_0\bar{b}$. Setting $\mathcal{M} = \mathcal{M}_0 = \mathcal{N}_0$, we can repeatedly apply Lemma 6.7 to construct an elementary chain (solid arrows indicate a proto-correspondence):



where both $\Sigma_i \subseteq \Pi_i^{-1} \subseteq \Sigma_{i+1}$ and $\mathcal{M}_i <^+ \mathcal{N}_i <^+ \mathcal{M}_{i+1}$ for each $i < \omega$. Define:

$$\begin{aligned} \mathcal{N} &= \bigcup_{i < \omega} \mathcal{N}_i = \bigcup_{i < \omega} \mathcal{M}_i \\ \Pi &= \bigcup_{i < \omega} \Pi_i \end{aligned}$$

It is routine to check that Π and \mathcal{N} have the required properties. \square

We can now obtain our \mathcal{L}^- -analogue of Theorem 6.2:

Theorem 6.9. For any \mathcal{L} -structure \mathcal{M} , the following are equivalent:

1. $\mathcal{M} \models \phi(\bar{a}) \leftrightarrow \phi(\bar{b})$, for all every $\phi \in \mathcal{L}_n^-$
2. There is an $\mathcal{N} >^+ \mathcal{M}$ and a relativity Π on \mathcal{N} such that $\bar{a}\Pi\bar{b}$
3. There is an $\mathcal{N} >^- \mathcal{M}$ and a relativity Π on \mathcal{N} such that $\bar{a}\Pi\bar{b}$

Proof. (1) \Rightarrow (2). Use Lemma 6.8 to obtain a near-correspondence, then use Lemma 4.9 to extend this to a relativity.

(2) \Rightarrow (3). Trivial.

(3) \Rightarrow (1). Π is a relativity, and hence a near-correspondence by Lemma 2.7; the result now follows since $\mathcal{M} <^- \mathcal{N}$. \square

This Theorem lends yet more weight to the claim that relativeness correspondences are the \mathcal{L}^- -analogue of symmetries. Moreover, it immediately yields a new way to characterise \approx_p and \approx_m (compare Lemma 6.3):

Lemma 6.10. For any \mathcal{L} -structure \mathcal{M} :

1. $a \approx_m b$ in \mathcal{M} iff there is an $\mathcal{N} >^+ \mathcal{M}$ in which $a r_m b$
2. $a \approx_p b$ in \mathcal{M} iff there is an $\mathcal{N} >^+ \mathcal{M}$ in which $a r_p b$

Moreover, both claims hold with $>^-$ in place of $>^+$.

Before continuing with the main aims of this paper, it is worth briefly stopping to smell the roses. Theorem 6.2 is sometimes used as a stepping stone to the following foundational result of definability theory (notation clarified in endnote):¹²

Theorem 6.11 (Beth–Svenonius Theorem, \mathcal{L}^+ -case). For any \mathcal{L} -structure \mathcal{M} with $R \notin \mathcal{L}$ and $U \subseteq M^n$, the following are equivalent:

1. $(\mathcal{M}, U) \models \forall \bar{v}(\phi(\bar{v}) \leftrightarrow R\bar{v})$ for some $\phi \in \mathcal{L}_n^+$.
2. For every $(\mathcal{N}, V) >^+ (\mathcal{M}, U)$ and every symmetry π of \mathcal{N} : $V = \pi(V)$.
3. For any \mathcal{L} -structure \mathcal{N} and any sets $V_0, V_1 \subseteq N^n$: if $(\mathcal{N}, V_0), (\mathcal{N}, V_1)$ and (\mathcal{M}, U) all satisfy the same \mathcal{L}^+ -sentences, then $V_0 = V_1$.

Pleasingly, we can use Theorem 6.9 as a stepping-stone to an \mathcal{L}^- -analogue of this result; indeed, the main steps are exactly as in the \mathcal{L}^+ -case.¹³

Theorem 6.12 (Beth–Svenonius Theorem, \mathcal{L}^+ -case). For any \mathcal{L} -structure \mathcal{M} with $R \notin \mathcal{L}$ and $U \subseteq M^n$, the following are equivalent:

1. $(\mathcal{M}, U) \models \forall \bar{v}(\phi(\bar{v}) \leftrightarrow R\bar{v})$ for some $\phi \in \mathcal{L}_n^-$.
2. For every $(\mathcal{N}, V) >^+ (\mathcal{M}, U)$ and every relativity Π of \mathcal{N} : $V = \Pi(V)$
3. For every $(\mathcal{N}, V) >^- (\mathcal{M}, U)$ and every relativity Π of \mathcal{N} : $V = \Pi(V)$
4. For any \mathcal{L} -structure \mathcal{N} and any sets $V_0, V_1 \subseteq N^n$: if $(\mathcal{N}, V_0), (\mathcal{N}, V_1)$ and (\mathcal{M}, U) all satisfy the same \mathcal{L}^- -sentences, then $V_0 = V_1$.

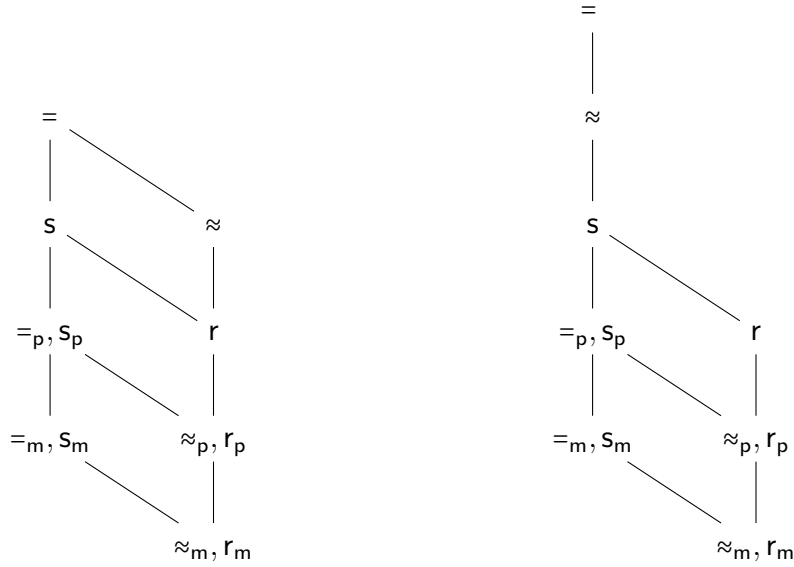
¹²Beth [1] proved (1) \Leftrightarrow (3); Svenonius [27] proved (1) \Leftrightarrow (2). (\mathcal{M}, U) is the $\mathcal{L} \cup \{R\}$ -structure formed from \mathcal{M} by allowing R to pick out U . As one would expect, $\pi(V) = \{\pi(\bar{e}) \mid \bar{e} \in V\}$, and equally $\Pi(V) = \{\bar{e} \mid \text{there are } \bar{d} \in V \text{ such that } \bar{d}\Pi\bar{e}\}$.

¹³The only difficult step in either Theorem is (2) \Rightarrow (1). For the \mathcal{L}^+ -case, see e.g. Poizat [23, Proposition 9.2]. To prove the \mathcal{L}^- -case, we simply tweak Poizat’s proof by invoking Theorem 6.9 rather than Theorem 6.2, and considering n -types in the sense of \mathcal{L}_n^- , rather than \mathcal{L}_n^+ . Dellunde [10, p. 5] and Keisler and Miller [15, p. 3] have shown that \mathcal{L}_n^- -types behave as one would hope.

7 Entailments in the finitary case

The results of the previous section immediately yield a special case of Theorem 3.3, obtained by restricting our attention to *finitary* structures.¹⁴ This special case has already attracted some attention, since it is philosophically interesting,¹⁵ and the following result completes the picture.

Theorem 7.1 (Entailments between the grades, finite structures). These Hasse Diagrams characterise the entailments between our grades of discrimination, when we restrict our attention to finite structures:



The left diagram is restricted to finite structures with arbitrary signatures; the right diagram is restricted to finite structures with *relational* signatures.

Proof. Most of this is supplied by Theorem 3.3. For the remainder observe that if \mathcal{M} is finite, then $\mathcal{N} >^+ \mathcal{M}$ iff $\mathcal{M} = \mathcal{N}$. It follows from Lemma 6.3 that \approx_p entails s_p and that \approx_m entails s_m ; and similarly with Lemma 6.10. \square

¹⁴An alternative route to Theorem 7.1 merits comment. We can use finitary isomorphism systems to prove the coincidence of grades of \mathcal{L}^+ -indiscernibility with grades of symmetry in finite structures, without invoking the results from §6. (For an introduction to finitary isomorphism systems, see Ebbinghaus et al. [11, chapter XI].) Casanovas et al. [7, Definitions 4.1–4.2] define the \mathcal{L}^- -analogue of finitary isomorphism systems. It turns out that we can use these, analogously, to prove the coincidence of grades of \mathcal{L}^- -indiscernibility with grades of relativity in finite structures.

¹⁵Caulton and Butterfield [8, Theorem 2] prove a special case of the mutual entailment between s_p and \approx_p , and s_m and \approx_m , on the assumption that \mathcal{L} is finite and relational. Ketland's [17, p. 2] attention is entirely restricted to finite relational signatures. Linnebo and Muller [19, Theorem 3] note that *witness-discernibility* (a further notion, which I have not discussed) is equivalent to \approx_p in finite structures, and outline several reasons for focussing on finite structures.

However, a little work will yield an even stronger result: the grades of relativity and the grades of \mathcal{L}^- -indiscernibility also entail each other when the structure's \approx -quotient is finite. To show this, we need a few results. The first tells us when \approx is definable in a structure.¹⁶

Lemma 7.2. Let \mathcal{M} be any \mathcal{L} -structure. If either \mathcal{L} is finite and relational or \underline{M} is finite, then there is an \mathcal{L}_2^- -formula which defines \approx in \mathcal{M} . However, the restrictions are necessary.

Proof. Case when \mathcal{L} is finite and relational. Let ϕ_1, \dots, ϕ_n enumerate all the atomic \mathcal{L} -formulas. By Lemma 2.2, the following \mathcal{L}_2^- -formula defines \approx in \mathcal{M} :

$$\bigwedge_{i=1}^n \forall \bar{v} (\phi_i(x, \bar{v}) \leftrightarrow \phi_i(y, \bar{v}))$$

Case when \underline{M} is finite. Let e_1, \dots, e_m exhaustively enumerate \underline{M} without repetition. So for all $i \neq j$ between 1 and m , we have $e_i \approx e_j$; hence by Lemma 2.2 there is some $\phi_{i,j} \in \mathcal{L}_2^-$ such that $\mathcal{M} \models \phi_{i,j}(e_i, e_i) \leftrightarrow \phi_{i,j}(e_i, e_j)$, and so the following \mathcal{L}_2^- -formula defines \approx in \mathcal{M} :

$$\bigwedge_{i \neq j} (\phi_{i,j}(x, x) \leftrightarrow \phi_{i,j}(x, y))$$

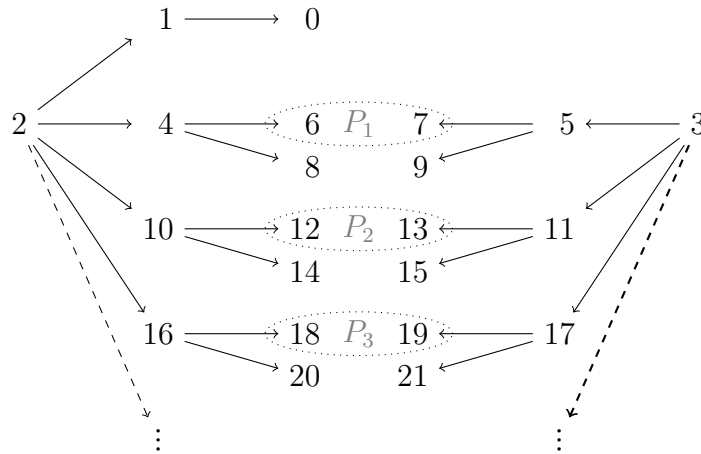
The necessity of the restrictions. Let \mathcal{L} contain one-place predicates P_i for all $i < \omega$, and a single two-place predicate R . Define:

$$H := \mathbb{N}$$

$$P_i^{\mathcal{H}} := \{6i, 6i + 1\}, \text{ for all } 0 < i < \omega$$

$$R^{\mathcal{H}} := \{\langle 2, 1 \rangle, \langle 1, 0 \rangle\} \cup \{\langle 2, 6n - 2 \rangle, \langle 6n - 2, 6n \rangle, \langle 6n - 2, 6n + 2 \rangle \mid 0 < n < \omega\} \cup \{\langle 3, 6n - 1 \rangle, \langle 6n - 1, 6n + 1 \rangle, \langle 6n - 1, 6n + 3 \rangle \mid 0 < n < \omega\}$$

We can represent \mathcal{H} more perspicuously as follows:



¹⁶For the case where \mathcal{L} is finite and relational, see Ketland [17, Definition 2.3].

I claim that $\underline{\mathcal{H}} \models \phi(\underline{2}) \leftrightarrow \phi(\underline{3})$ for all $\phi \in \mathcal{L}_1^-$. To prove this, fix $\phi \in \mathcal{L}_1^-$ and let \mathcal{K} be the (necessarily finite) set of \mathcal{L} -predicates appearing in ϕ . Where \mathcal{H}^* is the \mathcal{K} -reduct of \mathcal{H} , we have $\underline{2} \text{ s}_p \underline{3}$ in $\underline{\mathcal{H}}^*$, and hence $2 \text{ r}_p 3$ in \mathcal{H}^* by Lemma 4.10. Lemma 2.7 now yields that $\mathcal{H}^* \models \phi(2) \leftrightarrow \phi(3)$, and hence $\mathcal{H} \models \phi(2) \leftrightarrow \phi(3)$. Now apply Lemma 4.3.

However, where $\psi \in \mathcal{L}_1^+$ abbreviates:

$$\exists x(Rvx \wedge \forall y \forall z((Rxy \wedge Rxz) \rightarrow y = z)))$$

we have $\underline{\mathcal{H}} \models \psi(\underline{2}) \wedge \neg\psi(\underline{3})$. So no \mathcal{L}_2^- -formula can define $=$ in $\underline{\mathcal{H}}$; and hence no \mathcal{L}_2^- -formula can define \approx in \mathcal{H} , by Lemma 4.3. \square

We already knew that $a \approx b$ in \mathcal{M} iff $\underline{a} = \underline{b}$ in $\underline{\mathcal{M}}$. Lemma 7.2 allows us, under special circumstances, to obtain analogous results for our other grades of indiscernibility:

Lemma 7.3. Let \mathcal{M} be any \mathcal{L} -structure. If either \mathcal{L} is finite and relational or $\underline{\mathcal{M}}$ is finite, then:

1. $a \approx_p b$ in \mathcal{M} iff $\underline{a} =_p \underline{b}$ in $\underline{\mathcal{M}}$
2. $a \approx_m b$ in \mathcal{M} iff $\underline{a} =_m \underline{b}$ in $\underline{\mathcal{M}}$

Proof. By Lemma 7.2, given either assumption, some \mathcal{L}_2^- -formula defines \approx in \mathcal{M} . The same formula defines $=$ in $\underline{\mathcal{M}}$, and the result follows via Lemma 4.3. \square

Finally, the Galois Connection of §4 allows us to extend Theorem 7.1, as desired:

Lemma 7.4. For structures with finite \approx -quotients:

1. \approx_p entails r_p , and vice versa
2. \approx_m entails r_m , and vice versa.

Proof. Combine Theorem 7.1 with Lemmas 7.3 and 4.10. \square

Note that Theorem 7.1 does not hold for grades of \mathcal{L}^+ -discernibility/symmetry. To see this, let \mathcal{E}^* be a superstructure of \mathcal{E} obtained by making R reflexive. Whilst $\underline{\mathcal{E}}^*$ has only two members and its signature is finite and relational, no symmetry on \mathcal{E}^* sends 1 to any element in $\mathbb{R} \setminus \mathbb{N}$.

8 Capturing grades of discrimination

All twelve grades of discrimination have fairly straightforward definitions. However, the grades of indiscernibility are defined in terms of satisfaction of object-language formulas, whereas the grades of symmetry and relativity are defined metalinguistically. It is natural to ask whether this

is essential. More precisely, I shall ask which of the grades are *capturable*, in the following sense:

Definition 8.1. For any \mathcal{L} -structure \mathcal{M} , say that $\Gamma \subseteq \mathcal{L}_2$ *captures* R in \mathcal{M} iff:

$$\text{for all } a, b \in M : aRb \text{ in } \mathcal{M} \text{ iff } \mathcal{M} \models \phi(a, b) \text{ for every } \phi \in \Gamma$$

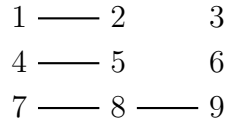
Say that R is *capturable⁺* in \mathcal{M} iff some $\Gamma \subseteq \mathcal{L}_2^+$ captures R in \mathcal{M} (similarly for *capturable⁻*). Say that R is *universally capturable⁺* iff some single $\Gamma \subseteq \mathcal{L}_2^+$ captures R in every \mathcal{L} -structure (similarly for *universally capturable⁻*).

I shall consider capturability for each of the three families of grades of discrimination, starting with the grades of indiscernibility:

- Lemma 8.2.**
1. $=, =_p, =_m$ are universally capturable⁺
 2. $\approx, \approx_p, \approx_m$ are universally capturable⁻
 3. There is a structure in which none of $=, =_p$ and $=_m$ is capturable⁻

Proof. (1) and (2). Obvious.

(3). Let \mathcal{I} be the following graph:



For all $\phi \in \mathcal{L}_2^-$, $\mathcal{I} \models \phi(1, 4) \leftrightarrow \phi(1, 7)$ even though $1 =_p 4$ and $1 \neq_m 7$. □

- Lemma 8.3.**
1. s is universally capturable⁺
 2. For finite structures: s_p and s_m are universally capturable⁺
 3. There is a structure in which s_p and s_m are not capturable⁺
 4. There is a finite structure in which none of s, s_p and s_m is capturable⁻

Proof. (1). For each atomic $\phi \in \mathcal{L}_{n+2}^+$, define:

$$x \simeq_\phi y := \forall \bar{v} \left(\bigwedge_{i=1}^n (v_i \neq x \wedge v_i \neq y) \rightarrow [\phi(x, y, \bar{v}) \leftrightarrow \phi(y, x, \bar{v})] \right)$$

By Lemma 2.4, s is universally captured⁺ by the set of all such \simeq_ϕ .

(2). From Theorem 7.1 and Lemma 8.2.

(3). Let \mathcal{J} comprise two disjoint copies of the complete countable graph, with a disjoint copy of a complete uncountable graph, i.e.:

$$\begin{aligned} J &:= \mathbb{R} \\ R^{\mathcal{J}} &:= \{ \langle m, n \rangle \in \mathbb{N}^2 \mid m \neq n \text{ and } m+n \text{ is even} \} \cup \{ \langle p, q \rangle \in (\mathbb{R} \setminus \mathbb{N})^2 \mid p \neq q \} \end{aligned}$$

By taking a Skolem Hull containing 1, 2 and some $e \in \mathbb{R} \setminus \mathbb{N}$, it is clear that:

$$\mathcal{J} \models \phi(1, 2) \leftrightarrow \phi(1, e)$$

for any $\phi \in \mathcal{L}_2^+$. However, $1 \text{ s}_p 2$ in \mathcal{J} , whereas $1 \not\text{ s}_m e$ in \mathcal{J} .

(4) In \mathcal{I} from Lemma 8.2, $1 \text{ s } 2$, whereas $7 \not\text{ s}_m 8$. However, for all $\phi \in \mathcal{L}_2^-$, $\underline{\mathcal{I}} \models \phi(\underline{1}, \underline{2}) \leftrightarrow \phi(\underline{7}, \underline{8})$ and hence $\mathcal{I} \models \phi(1, 2) \leftrightarrow \phi(7, 8)$. \square

Lemma 8.4. 1. r is universally capturable⁻

2. For structures with finite \approx -quotients: r_p and r_m are universally capturable⁻

3. There is a structure in which neither of r_p and r_m is capturable⁺

Proof. (1). Let Γ be the set of all \mathcal{L}_2^- -formulas of the form:

$$\forall \bar{v} \left(\bigwedge_{i=1}^n [\phi_i(x, x) \wedge \neg \phi_i(x, v_i) \wedge \psi_i(y, y) \wedge \neg \psi_i(y, v_i)] \rightarrow [\theta(x, y, \bar{v}) \leftrightarrow \theta(y, x, \bar{v})] \right)$$

for any $n < \omega$, any $\phi_1, \dots, \phi_n, \psi_1, \dots, \psi_n \in \mathcal{L}_2^-$, and any $\theta \in \mathcal{L}_{n+2}^-$. I claim that Γ captures r in any \mathcal{L} -structure \mathcal{M} .

First, suppose $a \text{ r } b$ in \mathcal{M} . Fix some $\gamma \in \Gamma$, and some $\bar{e} \in M^n$. Suppose that:

$$\mathcal{M} \models \bigwedge_{i=1}^n [\phi_i(a, a) \wedge \neg \phi_i(a, e_i) \wedge \psi_i(b, b) \wedge \neg \psi_i(b, e_i)]$$

Then by Lemma 2.2, $e_i \not\approx a$ and $e_i \not\approx b$ for each $1 \leq i \leq n$. Since $a \text{ r } b$, Lemma 2.7 tells us that $\mathcal{M} \models \theta(a, b, \bar{e}) \leftrightarrow \theta(b, a, \bar{e})$. Hence $\mathcal{M} \models \gamma(a, b)$, for any $\gamma \in \Gamma$.

Next, suppose $\mathcal{M} \models \gamma(a, b)$, for all $\gamma \in \Gamma$. I claim that the following is a near-correspondence from \mathcal{M} to \mathcal{M} :

$$\Pi = \{\langle a, b \rangle, \langle b, a \rangle\} \cup \{\langle x, x \rangle \mid x \not\approx a \text{ and } x \not\approx b\}$$

To show this, fix $n < \omega$, $\theta \in \mathcal{L}_{n+2}^-$ and $\bar{e} \in M^n$ such that $e_i \not\approx a$ and $e_i \not\approx b$ for each $1 \leq i \leq n$. Since each $e_i \not\approx a$ and $e_i \not\approx b$, by Lemma 2.2 there are formulas $\phi_i, \psi_i \in \mathcal{L}_2^-$ for each $1 \leq i \leq n$ such that $\mathcal{M} \models \phi_i(a, a) \wedge \neg \phi_i(a, e_i)$ and $\mathcal{M} \models \psi_i(b, b) \wedge \neg \psi_i(b, e_i)$. Conjoining these, we get:

$$\mathcal{M} \models \bigwedge_{i=1}^n [\phi_i(a, a) \wedge \neg \phi_i(a, e_i) \wedge \psi_i(b, b) \wedge \neg \psi_i(b, e_i)]$$

Since $\mathcal{M} \models \gamma(a, b)$ for all $\gamma \in \Gamma$, we obtain that, for all $\theta \in \mathcal{L}_{n+2}^-$:

$$\mathcal{M} \models \theta(a, b, \bar{e}) \leftrightarrow \theta(b, a, \bar{e})$$

Generalising, Π is a near-correspondence. By the Galois Connection of Theorem 4.8, $(\Pi^c)^e$ is a relativity on \mathcal{M} ; and so $a \text{ r } b$.

(2). From Lemmas 7.4 and 8.2.

(3). Exactly as in Lemma 8.3, case (3). \square

Lemmas 8.2–8.4 can be summarised as follows:

Theorem 8.5 (Capturing the grades). The following table exhaustively details the capturability of each grade of discrimination:

Grade	Capturable ⁺	Capturable [−]
=	✓	×
= _p	✓	×
= _m	✓	×
≈	✓	✓
≈ _p	✓	✓
≈ _m	✓	✓
s	✓	×
s _p	f	×
s _m	f	×
r	✓	✓
r _p	fq	fq
r _m	fq	fq

The table of Theorem 8.5 should be read with the following key:

- ✓ universally capturable
- × there is an \mathcal{L} -structure in which the grade is not capturable
- f** universally capturable when we restrict attention to finite structures; but there are counterexamples elsewhere
- fq** universally capturable when we restrict attention to structures with finite \approx -quotients; but there are counterexamples elsewhere

This demonstrates, once again, that grades of \mathcal{L}^+ -discernibility are to grades of symmetry, as grades of \mathcal{L}^- -discernibility are to grades of relativity. More interestingly, though, Theorem 8.5 bears directly upon the philosophical search for reductive criteria of identity.

As mentioned in §1, much of the interest in grades of discrimination comes from their potential to provide us with a criterion of identity, possibly a reductive one. However, if a grade of discrimination cannot be captured by some set of formulas in the object language, this should bar it from use in any reductive criterion of identity. After all, if the grade must be invoked as a *primitive* at the level of the object language, it is unclear why we should not simply allow ourselves to take identity itself as a primitive in the object language. The situation will be no better, in this regard, if the grade can only be captured⁺ and not captured[−]. Consequently, no grade of \mathcal{L}^+ -indiscernibility or symmetry can provide a *reductive* criterion of identity.

The remaining candidates for reductive criteria of identity are therefore the grades of \mathcal{L}^- -indiscernibility and relativity. However, in the special cases when they are capturable⁻—which we require if we seek a reductive criterion of identity—two of the grades of \mathcal{L}^- -indiscernibility are simply co-extensive with two of the grades of \mathcal{L}^- -indiscernibility (see Theorem 7.1). Hence the only plausible distinct candidates for a reductive criterion of identity are, in order of entailment: \approx , r , \approx_p and \approx_m .

This does not show, though, that the remaining grades of discrimination are philosophically uninteresting. After all, we might simply be interested in providing an illuminating but *non-reductive* answer to the general question: *When are objects identical?* To repeat an example from §1: if we have become convinced that nature abhors a (non-trivial) symmetry, then s_m could serve as a non-reductive, non-trivial criterion of identity, even though it is uncapturable⁺.

9 Symmetry in all elementary extensions

In §6, I connected the grades of indiscernibility with the existence of a symmetry/relativity in *some* elementary extensions. To close this paper, I wish to consider what happens when we consider the existence of a symmetry or relativity in *all* elementary extensions. In particular, I shall demonstrate a neat connection between \approx and symmetries in elementary extensions. To show this, I first require a general method for constructing such elementary extensions:¹⁷

Lemma 9.1. Let \mathcal{M} be an \mathcal{L} -structure with $a \in M$, and let D be a set such that $M \cap D = \emptyset$. Then there is an \mathcal{L} -structure $\mathcal{N} >^- \mathcal{M}$ with $N = M \cup D$, such that $a \approx d$ in \mathcal{N} for all $d \in D$.

Proof. Define $\sigma : N \longrightarrow M$ by: $\sigma(x) = x$ if $x \in M$, and $\sigma(d) = a$ if $d \in D$. Set:

$$\begin{aligned} R^{\mathcal{N}} &= \{\bar{e} \in N^n \mid \sigma(\bar{e}) \in R^{\mathcal{M}}\} && \text{all } n\text{-place } \mathcal{L}\text{-predicates } R \\ f^{\mathcal{N}}(\bar{e}) &= f^{\mathcal{M}}(\sigma(\bar{e})) && \text{all } n\text{-place } \mathcal{L}\text{-function-symbols } f \text{ and all } \bar{e} \in N^n \end{aligned}$$

I claim that, for each \mathcal{L} -term τ , all $\bar{d} \in D^m$ and all $\bar{e} \in M^n$:

$$\tau^{\mathcal{N}}(\bar{d}, \bar{e}) = \tau^{\mathcal{N}}(\bar{a}, \bar{e}) = \tau^{\mathcal{M}}(\bar{a}, \bar{e})$$

(where $a_i = a$ for all $1 \leq i \leq m$). This is proved by induction on complexity. The case where τ is an \mathcal{L} -function symbol is given. Now suppose the claim

¹⁷Monk [21, Theorem 29.16] described this explicitly; Grzegorzcyk [12, p. 41] earlier mentioned it in passing, implying it was mathematical folklore. The method was rediscovered by philosophers, e.g. Ketland [17, p. 7] and Ladyman et al. [18, Theorem 8.14]. However, all these authors restrict their attentions to relational signatures.

holds for τ_1, \dots, τ_k and consider $\tau(\bar{x}, \bar{y}) = f(\tau_1(\bar{x}, \bar{y}), \dots, \tau_k(\bar{x}, \bar{y}))$. Then:

$$\begin{aligned}\tau^{\mathcal{N}}(\bar{d}, \bar{e}) &= f^{\mathcal{N}}(\tau_1^{\mathcal{N}}(\bar{d}, \bar{e}), \dots, \tau_k^{\mathcal{N}}(\bar{d}, \bar{e})) \\ &= f^{\mathcal{N}}(\tau_1^{\mathcal{N}}(\bar{a}, \bar{e}), \dots, \tau_k^{\mathcal{N}}(\bar{a}, \bar{e})) = \tau^{\mathcal{N}}(\bar{a}, \bar{e}) \\ &= f^{\mathcal{N}}(\tau_1^{\mathcal{M}}(\bar{a}, \bar{e}), \dots, \tau_k^{\mathcal{M}}(\bar{a}, \bar{e})) \\ &= f^{\mathcal{M}}(\tau_1^{\mathcal{M}}(\bar{a}, \bar{e}), \dots, \tau_k^{\mathcal{M}}(\bar{a}, \bar{e})) = \tau^{\mathcal{M}}(\bar{a}, \bar{e})\end{aligned}$$

This proves the claim. Hence, for all atomic $\phi \in \mathcal{L}_{m+n}^-$, all $\bar{d} \in D^m$ and all $\bar{e} \in M^n$:

$$\mathcal{N} \models \phi(\bar{d}, \bar{e}) \text{ iff } \mathcal{N} \models \phi(\bar{a}, \bar{e}) \text{ iff } \mathcal{M} \models \phi(\bar{a}, \bar{e})$$

So for all $d \in D$ we have $a \approx d$ in \mathcal{N} by Lemma 2.2; and moreover $\mathcal{N} >^- \mathcal{M}$. \square

Thus armed, I can connect \approx with symmetry in elementary extensions:

Lemma 9.2. For any \mathcal{L} -structure \mathcal{M} : if $a \mathbf{s}_m b$ in every $\mathcal{N} >^- \mathcal{M}$, then $a \approx b$ in \mathcal{M} .

Proof. Suppose $a \mathbf{s}_m b$ in every $\mathcal{N} >^- \mathcal{M}$. Let D be such that $M \cap D = \emptyset$ and $|D| > |\underline{b}_{\mathcal{M}}|$. Construct \mathcal{N} as in Lemma 9.1, so that $a \approx d$ in \mathcal{N} for all $d \in D$. Since $\mathcal{N} >^- \mathcal{M}$, by assumption there is a symmetry π on \mathcal{N} such that $\pi(a) = b$. So, for every $\phi \in \mathcal{L}_2^-$, and all $d \in D$, by Lemma 2.4:

$$\mathcal{N} \models \phi(b, b) \text{ iff } \mathcal{N} \models \phi(a, a) \text{ iff } \mathcal{N} \models \phi(a, d) \text{ iff } \mathcal{N} \models \phi(b, \pi(d))$$

Hence $\pi(d) \in \underline{b}_{\mathcal{N}}$ for every $d \in D$, by Lemma 2.2. Since π is a bijection, $|D| = |\{\pi(d) \mid d \in D\}| \leq |\underline{b}_{\mathcal{N}}|$. If $a \not\approx b$ in \mathcal{N} , then $|\underline{b}_{\mathcal{N}}| = |\underline{b}_{\mathcal{M}}|$, contradicting our choice of D ; so $a \approx b$ in \mathcal{N} . Since $\mathcal{M} <^- \mathcal{N}$ and \approx is universally capturable by Lemma 8.2, $a \approx b$ in \mathcal{M} . \square

Theorem 3.3 (left-diagram) entails us that there is no converse to 9.2 in the general case. However, we do obtain a converse in restricted circumstances:

Lemma 9.3. When \mathcal{L} is relational, for any \mathcal{L} -structure \mathcal{M} : $a \approx b$ in \mathcal{M} iff $a \mathbf{s}_m b$ in every $\mathcal{N} >^- \mathcal{M}$.

Proof. Immediate from Theorem 3.3 and Lemmas 8.2 and 9.2. \square

Moreover, we can strengthen Lemma 9.2 in the case of \mathbf{s} .

Lemma 9.4. Let \mathcal{M} be an \mathcal{L} -structure with $a \in M$ and $e \notin M$. Let $\mathcal{N} >^- \mathcal{M}$ be constructed as in Lemma 9.1, so that $N = M \cup \{e\}$ and $a \approx e$ in \mathcal{N} . If $a \mathbf{s} b$ in \mathcal{N} , then $a \approx b$ in \mathcal{M} .

Proof. Suppose $a \text{ s } b$ in \mathcal{N} , i.e. $\pi(a) = b$, $\pi(b) = a$, and $\pi(x) = x$ for all $x \notin \{a, b\}$ is a symmetry on \mathcal{N} . In particular, $\pi(e) = e$. Hence, invoking Lemma 2.4, for all $\phi \in \mathcal{L}_2^-$: $\mathcal{M} \models \phi(a, a)$ iff $\mathcal{N} \models \phi(a, a)$ iff $\mathcal{N} \models \phi(e, a)$ iff $\mathcal{N} \models \phi(e, b)$ iff $\mathcal{N} \models \phi(a, b)$ iff $\mathcal{M} \models \phi(a, b)$. Hence $a \approx b$ in \mathcal{M} by Lemma 2.2. \square

However, this strengthening of Lemma 9.2 is limited to the case of s . To see this, let \mathcal{K} comprise two disjoint copies of the complete countable graph, i.e.:

$$K = \mathbb{N}$$

$$R^{\mathcal{K}} = \{\langle m, n \rangle \in \mathbb{N}^2 \mid m \neq n \text{ and } m + n \text{ is even}\}$$

Whilst $1 \not\approx 2$ in \mathcal{K} , we can use Lemma 9.1 to add a single new element, e , such that $1 \approx e$, without disrupting the fact that $1 \text{ s }_p 2$. Moreover, nothing like Lemma 9.2 holds for relativities: Lemma 8.4 tells us that $a \text{ r } b$ in \mathcal{M} iff $a \text{ r } b$ in all $\mathcal{N} \succ^- \mathcal{M}$.

The results of §6 exhaustively detailed the connections between grades of discernibility and the existence of a symmetry/relativity in *some* elementary extension. The results of this section now exhaustively detail the connections between grades of discernibility and the existence of a symmetry/relativity in *all* elementary extensions. We thus have complete answers to several natural questions concerning the connection between grades of discrimination and elementary extensions.

10 Concluding remarks

Several recent technical-cum-philosophical papers have explored some of the grades of discrimination. This paper has pressed forward that technical investigation in many ways. To close, I shall emphasise two.

First, I have introduced *grades of relativity* to the philosophical literature—along with the notion of a near-correspondence, a relativeness correspondence, and a partial relativeness correspondence—and shown that these are the natural \mathcal{L}^- -analogues of the grades of symmetry.

Second, I have offered complete answers to the natural questions that arise concerning all twelve grades of discrimination. Indeed, the technical investigation of the grades of discrimination now seems to be complete.¹⁸

¹⁸Huge thanks to Øystein Linnebo and Sean Walsh, whose questions provided much of the original motivation for this paper, and whose subsequent comments were very helpful. Further thanks to Denis Bonnay, Adam Caulton, Fredrik Engström, Jeffrey Ketland, Richard Pettigrew, a referee for *RSL*, and a referee for this journal.

References

- [1] Beth, E. W., “On Padoa’s method in the theory of definition,” *Nederlandse Akademie van Wetenschappen. Proceedings. Series A. Indagationes Mathematicae*, (1953), pp. 330–9.
- [2] Black, M., “The identity of indiscernibles,” *Mind*, vol. 61 (1952), pp. 153–64.
- [3] Blok, W. J., and D. Pigozzi, “Protoalgebraic logics,” *Studia Logica*, vol. 45 (1986), pp. 337–69.
- [4] Bonnay, D., and F. Engström, “Invariance and definability, with and without equality,” arXiv:1308.1565 [math.LO], (2013).
- [5] Button, T., “Realistic structuralism’s identity crisis: A hybrid solution,” *Analysis*, vol. 66 (2006), pp. 216–22.
- [6] Button, T., *The Limits of Realism*, Oxford University Press, Oxford, 2013.
- [7] Casanovas, E., P. Dellunde, and R. Jansana, “On elementary equivalence for equality-free logic,” *Notre Dame Journal of Formal Logic*, vol. 37 (1996), pp. 506–22.
- [8] Caulton, A., and J. Butterfield, “On kinds of indiscernibility in logic and metaphysics,” *British Journal for the Philosophy of Science*, vol. 63 (2012), pp. 27–84.
- [9] Dellunde, P., “Equality-free logic: The method of diagrams and preservation theorems,” *Logic Journal of the IGPL*, vol. 7 (1999), pp. 717–32.
- [10] Dellunde, P., “Equality-free saturated models,” *Reports on Mathematical Logic*, vol. 37 (2003), pp. 3–22.
- [11] Ebbinghaus, H.-D., J. Flum, and W. Thomas, *Mathematical Logic*, Springer-Verlag, New York, 1984.
- [12] Grzegorzczuk, A., “On the concept of categoricity,” *Studia Logica*, vol. 13 (1962), pp. 39–65.
- [13] Hawley, K., “Identity and indiscernibility,” *Mind*, vol. 118 (2009), pp. 101–19.
- [14] Hilbert, D., and P. Bernays, *Grundlagen der Mathematik*, volume 1, Springer, Berlin, 1934.
- [15] Keisler, H. J., and A. W. Miller, “Categoricity without equality,” (MS). URL <http://www.math.wisc.edu/~keisler/equal.pdf>.

- [16] Ketland, J., “Structuralism and the identity of indiscernibles,” *Analysis*, vol. 66 (2006), pp. 303–15.
- [17] Ketland, J., “Identity and indiscernibility,” *The Review of Symbolic Logic*, vol. 4 (2011), pp. 171–85.
- [18] Ladyman, J., O. Linnebo, and R. Pettigrew, “Identity and discernibility in philosophy and logic,” *Review of Symbolic Logic*, vol. 5 (2012), pp. 162–86.
- [19] Linnebo, O., and F. A. Muller, “On witness-discernibility of elementary particles,” *Erkenntnis*, vol. 78 (2013), pp. 1133–42.
- [20] Marker, D., *Model Theory: An Introduction*, Springer, New York, 2002.
- [21] Monk, J. D., *Mathematical Logic*, Springer-Verlag, New York, 1976.
- [22] Muller, F. A., “The rise of relationals,” *Mind*, vol. 124 (2015), pp. 201–37.
- [23] Poizat, B., *A Course in Model Theory: An Introduction to Contemporary Mathematical Logic*, Springer-Verlag, New York, 2000.
- [24] Quine, W. v. O., *Word and Object*, MIT Press, Cambridge MA, 1960.
- [25] Quine, W. v. O., “Grades of discriminability,” *Journal of Philosophy*, vol. 73 (1976), pp. 113–6.
- [26] Shapiro, S., “An ‘*i*’ for an *i*: Singular terms, uniqueness, and reference,” *The Review of Symbolic Logic*, vol. 5 (2012), pp. 380–415.
- [27] Svenonius, L., “A theorem on permutation in models,” *Theoria*, vol. 25 (1959), pp. 173–8.
- [28] Waszkiewicz, J., “The notions of isomorphism and identity for many-valued relational structures,” *Studia Logica*, vol. 27 (1971), pp. 93–8.